Switched Systems as a Common Computational Principle: Connections Between Control Theory, Neural Networks and Context-Selected Linear Mappings

Abstract

Many branches of mathematics and engineering have independently developed theories based upon switching between relatively simple computational or dynamical components. These include switched control systems, hybrid dynamical systems, piecewise-linear systems, neural ordinary differential equations and ReLU neural networks. Although each field has developed its own terminology and mathematical tools, they appear to share a common computational principle: complex global behaviour emerging from context-dependent selection among simple local models. This note summarizes these connections and highlights areas where ideas from one discipline may inform another.

The Central Observation

Across multiple disciplines we repeatedly encounter the same architecture:

- many relatively simple local models,
- a mechanism that selects between them,
- global complexity emerging from composition rather than from highly complex local computation.

The mathematical machinery often differs.

The underlying computational idea appears remarkably similar.

Classical Switched Systems

Substantial existing literature

Typical keywords:

- switched systems
- switched linear systems
- switched affine systems
- switched dynamical systems
- mode switching
- supervisory control
- gain scheduling
- multiple-model control

These systems typically consist of several linear or mildly nonlinear dynamical models.

A switching mechanism determines which model is currently active.

Much of the research concerns

- stability,
- robustness,
- switching strategies,
- Lyapunov functions,
- dwell-time analysis.

Hybrid Systems

Extensive existing theory

Keywords:

- hybrid systems
- hybrid automata
- piecewise-smooth systems
- state machines
- finite automata
- event-triggered systems

These combine

- continuous dynamics,
- discrete state transitions.

This framework appears naturally suited to context-dependent computation.

Piecewise Linear Systems

Large mathematical literature

Keywords:

- piecewise linear systems
- piecewise affine systems
- piecewise smooth dynamics
- polyhedral partitions
- complementarity systems

These systems already study many of the mathematical properties that later appeared within ReLU networks.

Neural Networks

Modern ReLU networks introduce another member of this family.

Keywords:

- piecewise linear neural networks
- activation regions
- linear regions
- decision regions
- expressivity
- Jacobian analysis
- Lipschitz continuity

Within each activation region

- computation is linear,
- only the selected linear mapping changes.

Most theoretical work has concentrated on

- approximation power,
- optimization,
- expressivity,
- counting linear regions.

Relatively less attention appears to have been given to interpreting the network as a switched computational system closely related to classical control theory.

Neural Ordinary Differential Equations

Recent work has connected deep learning with continuous dynamical systems.

Keywords:

- Neural ODEs
- continuous-depth networks
- flow-based models
- dynamical systems
- differential equations
- invertible neural networks

These replace discrete layers with continuous flows.

The emphasis shifts from layer composition to trajectory evolution.

Context-Selected Linear Mappings

CCSLM suggests making the switching mechanism explicit.

Separate

- the data pathway,
- the context pathway.

The context determines which linear mapping is active.

This resembles supervisory control more closely than conventional neural network descriptions.

Topology

Several mathematical communities already investigate geometric structure.

Useful keywords:

- differential topology
- manifold learning
- diffeomorphisms
- homeomorphisms
- geometric deep learning
- bi-Lipschitz mappings
- topology preservation
- invertible neural networks

These provide mathematical tools for studying how computational systems deform geometric structure.

Optimization

Backpropagation introduces another viewpoint.

Each local linear mapping becomes

- a learned average,
- statistically regularized,
- adapted to nearby samples.

This raises an interesting question.

Does gradient descent implicitly encourage neighbouring local models to become geometrically compatible?

This question appears to connect

- optimization,
- topology,
- switched systems,
- statistical learning.

Robotics

Robotics has long relied upon switching between specialised controllers.

Keywords:

- behavioural robotics
- behaviour arbitration
- hierarchical control
- subsumption architecture
- motion primitives
- locomotion primitives
- finite-state locomotion
- hybrid control

Again, relatively simple local controllers are combined into sophisticated global behaviour.

A Possible Common Computational Principle

These fields may all be studying variations of the same computational architecture.

Area	Local primitive	Selection mechanism
Switched control	Linear dynamics	Supervisory controller
Hybrid systems	Continuous dynamics	Discrete state
Robotics	Motion/controller primitives	Behaviour arbitration
ReLU neural networks	Linear mappings	Activation pattern
Neural ODEs	Vector fields	Continuous evolution
CCSLM	Linear mappings	Explicit context

The mathematical languages differ.

The computational architecture appears remarkably similar.

An Open Question

The interesting research question may therefore not be

> "Can switched systems perform intelligent computation?"

but rather

> **To what extent are modern neural networks, switched control systems, hybrid systems and context-selected computation different mathematical descriptions of the same underlying computational principle?**

If this perspective proves useful, decades of work in switched systems, hybrid control and dynamical systems may become directly relevant to understanding the geometry, topology and optimisation of modern neural architectures.

Rather than replacing existing theory, the goal would be to connect mature bodies of knowledge that have largely developed in parallel using different terminology.